

创建与使用型人以优化机构知识库
——基于文献计量的实证研究

□顾立平

摘要 分析了香港城市大学经济与金融学系的博硕士论文，指出作者们有不同的特征和习惯，因而需要不同程度的建议和信息服务。结论是采用书目计量的方法分析机构知识库，是创建与使用型人（多重用户模型）来扩展知识服务的新取向之一，有助于提高机构知识库的使用率。

关键词 机构知识库 文献计量 知识服务 用户行为 用户模型

1 研究背景与目的

英国伦敦大学的用户行为观察研究[1]对图书馆开展新一轮知识服务颇有启发，即：透过计算主题类别的集中度与整理核心期刊的影响度，而对经济与商业类研究生，进行文献与知识推荐的服务试点。此例若成，估计未来会推展到不同学科与不同程度的用户群体中去。除乐观其成外，本文也从机构知识库的角度，采用文献计量的方法、型人模型的理论去拓展它的研究范畴与深度。

机构知识库（IR）目前面临许多共同难题，例如：内容尚少、效果不佳、任务受到质疑、技术有待更新、缺少真正意义上的机构参与等。造成这些难题的原因很多，其中之一是"大多数教职员工仍然不明白自己为什么要使用机构知识库：它既不利于传播工作成果，也不利于得到经费资助。当然，也不会对资历和晋升有所帮助"[2][3]。换言之，增加用户参与感，或者利用知识服务使之感到机构知识库能够协助他们进行科学研究，是一项待开发的图书馆新型服务。

Maness等人较早进行这种新型服务的研究[4]，即透过改善交互设计而达到提升用户体验的IR型人研究。因为初期投入IR的图书馆面临许多未知变量，特别是Open Access以及IR的定位都尚未形成共识的时候，只能着重于数据库建设与呈缴制度的规范性研究。若要突破现有收藏与服务的局限，就需要开发新的服务，而具备实证性研究特点的型人就是这样一种可以拓展IR服务的模式。然而Maness等人尚未将IR的力量充分运用，对型人的使用也未脱离经典教科书[5]上的技术与范围，而缺乏独自开辟一条新方向的尝试，实为可惜。

综上，为扩展推荐信息的能力、提高机构知识库的质量以及利用型人创新图书馆服务，本文展开综合三者的实证研究设计。

2 研究设计

将IR、型人和用户行为结合是目的，手段是文献计量，实证对象是香港城市大学的经济与金融学院。

2.1 理论依据与研究框架

采用源自拉丁文的Personas作为理论依据。本文翻译为"型人"。尚存于世界的通用语言中，西班牙语仍使用Persona此字，其含义等于英语Person的含义，而Personas为西班牙语"人"这个名词的复数。开发Visual Basic的Alan Cooper曾说："型人不是真正的人，…"[6]。指称Personas在他的语言中不代表人，而是具有科学与工程的含义。反之，中文翻译Personas不应只照英文，而不参照拉丁文和西班牙文的原意。

国内对Personas的普遍翻译有二：文学理论领域翻译为"人物"，计算机领域翻译为"人物角色"，其定义略有不同。文学的Personas指虚构人物在作品中的应用与对读者的影响，而计算机中的Personas有两个意义，一是与"界面"有关，比如手机操作系统的模式，一是与"软件设计"有关，指程序员如何思考使用者行为。但图书情报学应用此字时，兼具上述所有含义。

故此，取中文里"典型"、"类型"、"模型"三字字

尾,结合中华儒家文化的"天(不可知)、地(可知)、人(可为)"概念中的"人"字为要义,在扩展 Personas 含义的基础上,翻译 Personas为"型人",应该较"人物"或"人物角色"更为简练和贴近英文、西班牙文与拉丁文的扩展义和原始义。

根据此理论而推导之研究框架,着重在"创建"与"使用"型人:创建型人着重在经验研究的基础上,依据实际情况或代换研究方法,或开发研究技术;使用型人着重在创建之后的实际应用,乃至于对基础技术的改进要求。故两者实为理论与实践之结合。

2 2 研究假设

根据行为主义与认知心理学的主要观点与假设,本文设定的研究假设为:(1) IR 中的文献的利用率与用户的特征和行为必然有关;(2)不同用户的行为需要不同的信息服务;(3)合适的信息服务会强化假设(1)的关联性。

2 3 研究问题

根据理论框架与研究假设,所研究问题包括三个主要问题:(1)用户的特征与行为;(2)从用户的特征与行为中,可分析出哪些不同类型的用户;(3)针对不同的用户,应该推荐哪些信息。

2 4 研究定位

采用书目计量法,但研究定位是案例研究(香港城市大学 – 经济与金融学系 – 收录于 IR 的博硕士论文),分析单位是单篇论文而非个人或机构。

另,数据源(http //dspace cityu edu hk/handle/2031/709/browse-title)是支持 Open Access的 IR 网址。最后访问日期是 2009年 10月 10日,后续研究人员可重复验证。

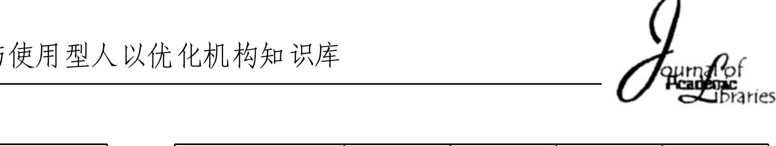
3 结果

3 1 用户特征与行为

据分析,30个样本的参考文献虽不尽相同,但集中反映出期刊所占的比率较高,其次是工作报告、会议论文、公司报表等具有研究价值(档案分析法)的灰色文献;一般列举少量 Cambridge或 Harvard出版社的专著作为理论框架或实验技术的依据,个别作者(如:Chong B-P-S和 Law K-C)偏好大量图书与期刊。一般不采用网络文献,仅部分作者(如:Lo W-L和 Chong B-P-S)依需要酌量参考,见表 1。

表 1 作者使用的文献类别

时间	作者	图书	期刊	网络	其他	大比例偏重的数据来源
1994	Shum, C-K-R	3	90	0	8	Journal of Finance
2001	Lo, W-L	4	24	18	16	Annual (different companies)
2003	Law, K-C	22	40	0	63	Working paper
2003	Zhou, K-G	1	67	0	4	Average
2004	Hung Y-K	7	100	6	9	Journal of Accounting Research Harvard Business Review Journal of Finance
2004	Sun, Y-J	17	132	0	13	American Economic Review
2004	Lam, E-C-F-Y	1	41	1	4	Journal of Political Economy Econometrica
2004	Jing L-H	7	90	0	14	Journal of Financial Economics Journal of Finance
2004	Yao, J	16	92	0	12	Journal of Financial Economics Journal of Financial Intermediation
2005	Cai, J	15	57	0	11	Average
2005	Ho, W-H	3	16	0	3	Average
2005	Wu, L-L	8	125	0	2	Journal of Financial Economics
2005	Chong B-P-S	40	131	12	25	Conferences paper
2005	Li, Y	11	55	9	19	Internet Engineering Task Force
2006	Kwok, W-H	17	111	0	20	Average
2006	Feng L-L	3	19	0	5	Average
2006	Guo, F	2	64	0	2	Journal of Finance Journal of Economics
2006	Jor, W-S	5	10	0	5	Average
2006	Ge, H	6	36	0	17	Average
2006	Zhou W-Q	6	57	2	24	Journal of Financial Economics
2006	Yip M-W	15	49	0	17	Average
2007	Hong X-P	1	31	0	6	Journal of Finance
2007	Chen, Q	2	102	0	15	Average
2007	Qi, Y-H	4	73	8	4	Journal of Financial Economics Journal of Finance
2007	Au Y-C-K	3	68	0	6	Journal of Financial Economics Journal of Finance



时间	作者	图书	期刊	网络	其他	大比例偏重的数据来源
2007	Lui, C-W-K	2	113	0	0	Journal of Financial Economics Journal of Finance Financial Management
2007	Zhang A	5	43	0	21	Journal of International Economics
2007	Ho, C-Y	8	45	0	18	Journal of Political Economy
2007	Chow, F-W-M	0	28	1	9	Financial Analysis Journal
2008	Wong L-K	1	21	1	2	Journal of Financial Economics

进一步分析,可以发现多数引用的期刊是: *Journal of Political Economy*、*American Economic Review*、*Journal of Financial Economics*, *Journal of Finance*, *Econometrica*, *Harvard Business Review*、*Financial Analysis Journal*。

3.2 用户的类型

以图书馆员对"分类"的敏锐直觉,可以观察到:作者虽然同属经济与金融专业,但所关心的区域经济问题有所不同,见表 2。

表 2 作者关心的区域经济问题

作者	中国香港	中国大陆	其他地区	不涉地区
Shum, C-K-R	●			
Lo, W-L	●			
Law, K-C				●理论
Zhou, K-G	●			
Hung, Y-K	●			
Sun, Y-J			●马来西亚	
Lam, E-C-F-Y				●理论
Jing, L-H	●	●		
Yao, J			●日本	
Cai, J	●			
Ho, W-H				●知识产权
Wu, L-L			●日本	
Chong, B-P-S	●			●电子商务
Li, Y				●技术
Kwok, W-H		●		
Feng, L-L		●		
Guo, F			●日本	
Jor, W-S	●	●		
Ge, H				●行贿问题
Zhou, W-Q	●			
Yip, M-W		●		
Hong, X-P	●			
Chen, Q			●日本	
Qi, Y-H	●			
Au, Y-C-K				●资本结构
Lui, C-W-K			●澳大利亚	
Zhang, A			●欧美各国	
Ho, C-Y		●		
Chow, F-W-M	●		●英国	
Wong, L-K	●	●		

因研究样本来自香港城市大学,因此偏好香港并不为其奇,但有越来越偏向于研究中国大陆的经济问题的倾向。此外,以日本为首的其他地区经济研究(在经济学中不称"案例研究",也非图书馆分类的"比较经济",而是以企业研究或制度研究等名义称之,但不见于关键词和题名)也是另一特征。而理论研究(指"以数学公式建立模型并用数据进行检验而无法证伪"的假说的集合)逐渐减少,但数理化模型则在"非理论"的研究中占越来越大的比例(篇幅和字数)。

3.3 IR 可提供的信息种类

根据现有 IR 内的信息,可发现不同学科、类别的信息,同时开放存取的信息组织机制使得多数 IR 能够互用其元数据内容,并且在开放存取的共同标准下,可实现多种信息的分类、收集与推荐,见表 3。

表 3 跨学科、类别、机构的 IR 信息种类

信息种类	作用	例如
跨学科的信息	以进行交叉学科研究之用	经济、管理、工程、情报、生物等
跨类别的信息	新想法、交流机会、数据、成果	期刊、会议、专著、文集、专著或文集的章节、研究报告、演示报告等
跨机构的信息	不同区域、不同观点、不同重点	香港大学、北京大学、台湾大学等

比方说,如果香港城市大学的金融系的研究生要撰写毕业论文,主题是有关科学园区的群聚效应,图书馆除了透过书目系统,推荐克鲁曼(Paul Robin Krugman)在麻省理工大学出版社出版的的专著或电子书外,也可通过 IR 推荐台湾政治大学吴思华的期刊论文等。



4 讨论

以 Chong B-P-S 研究电子商务为例(注:比方在淘宝网或阿里巴巴如何叫价与交易),该研究需要参考微观经济学、企业管理案例、传播学期刊文献和最新发展的会议论文。因此她的参考文献经计算呈现"Average"分布,而与 Li Y 不同。后者的研究,除了参考经济与金融期刊外,还大量参考了 *Internet Engineering Task Force* 等工程类别期刊的信息,经计算,其"非本专业"信息量占参考总量的一半。而 Wu L-L 的参考文献就几乎全部集中在经济与金融类期刊,而且明显以 *Journal of Financial Economics* 为主。见表 1。

换言之,关键在:图书馆能否加深知识服务的力度?

进一步分析这个关键,就问:图书馆能否对不同用户进行不同程度的知识服务?再进一步,就问:图书馆有没有条件?再进一步,再问:如果 IR 是唯一的补充条件,那么如何做?案例在哪?成功率有多大?再进一步,总结关键问题和子问题,一言以蔽之,就问:图书馆如何利用 IR 增加它的知识服务深度?

假设上述 30 个样本的作者,未来仍然持续相关研究,而需要图书馆的知识服务,那么图书馆在分析他们提交给 IR 的论文的基础上,可以推荐符合他们研究行为与特征的信息。这种"基于型人的知识服务机制",将是表 1 与表 2 的综合,见图 1。

根据图 1,分述知识服务机制如下:(1)为群 1 推荐期刊,但是区分全文获取和文摘获取,并利用关键词本体以及基于期刊网的发展,给予情报分析报告;(2)为群 2 推荐网络信息、会议信息、工作报告等"非正式"学术交流信息,但是需要经过整理,可以是网络导航也可以是内容聚合;(3)为群 3 推荐灰色文献,如实验数据、工作报告、公司报告等,以及"非专业"的期刊文献;(4)为群 4 推荐各种功能的选择,同时遴选经典图书,以及著名作者(例如诺贝尔奖得主)、知名出版社(如百年老店 Springer)、或者全球各地专业类极强的出版机关(如中国高等教育出版社)等,满足其心理与实际工作的需求。

建议图书馆员在讨论如何服务上述用户时,可在书目计量的基础上,通过访谈、照相(如果对方愿意)、咨询等方式创建与使用型人,以达到相互交流与累积经验以提升机构知识库的目标。

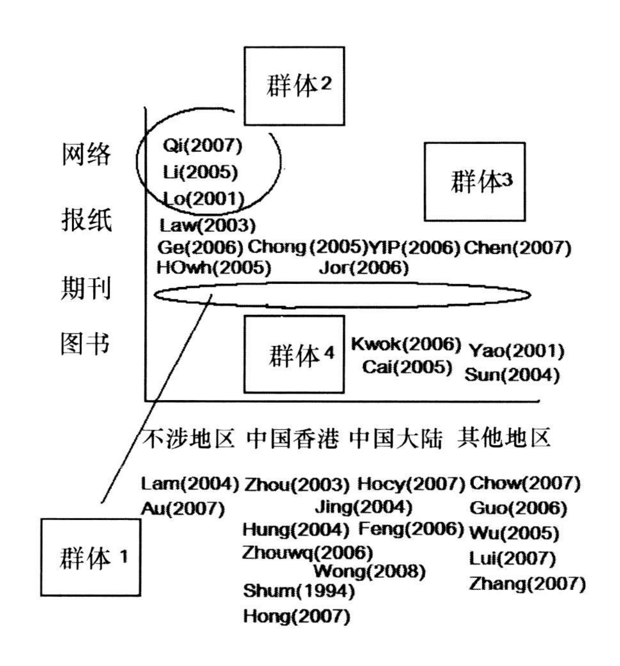


图 1 根据文献计量所划分之 IR 群组

参考文献

1 CIEBR University College London JISC User Behaviour Observational Study: Working paper on scholarly digital use and information seeking behaviour in business and economics. (2009-07-01) [2009-09-22]. http://www.jisc.ac.uk/whatwedo/programmes/inf11/userbehaviousagedata.aspx

2 Albanese A R. Institutional Repositories Thinking Beyond the Box-Repositories leapt into the national spotlight in 2008 Now what? (2009-04-02) [2009-09-09]

3 Albanese A R 著,熊瑞编译. 机构知识库:跳出"盒子"来思考——2008 年机构知识库成为国内关注焦点,而如今呢?. 图书情报工作动态, 2009(5): 17-18

4 Maness J M, Miaskiewicz T, Sumner T. Using Personas to Understand the Needs and Goals of Institutional Repository Users. D-Lib Magazine, 2008(9/10): 14 [2008-10-13]. http://www.dlib.org/dlib/september08/maness/09maness.html

5 Mulder S & Yaar Z. The user is always right: a practical guide to creating and using personas for the Web. Berkeley, CA: New Riders Press, 2006

6 Cooper A. The inmates are running the asylum: why high tech products drive us crazy and how to restore the sanity. USA, Sam publishing 2006: 0-1

作者单位:香港城市大学图书馆,香港

收稿日期:2009 年 11 月 30 日

(转第 112 页)

Survey and Analysis on Current Status of the Lecturers Teaching Literature Retrieval Course in Beijing Universities

Zhou Xuan Wang Mengli He Jia Du Weichun

Abstract A questionnaire was conducted on the Literature Retrieval Course lecturers at 33 colleges and universities in Beijing. The survey includes the integral structure of lecturers, their working situation, teaching management etc. The feedback result shows that lots of lecturers own a higher education degree, and the learning-origin structure is comparatively good. However, there still exits some dissatisfactory aspects. The disciplinary structure of lecturers was still to be improved, and the teaching skills to be enhanced. Also, the situation that lecturers lack training opportunists and assessment is a widespread and serious problem.

Keywords Information Literacy Education; Literature Retrieval Courses; Lecturers; Survey; Analysis

(接第 52页)

Creating and Using Personas for Improving Institutional Repository: An Empirical Study of Bibliometrics

Ku Liping

Abstract For expending the research field of "JISC User Behavior Observatoinal Study" report, adding the functions of institutional repository (IR) and innovating library services, the dissertations of economic and business colleage in the City University of Hong Kong were analyzed. The result is that there are different authors who need different levels of suggestions and information services. The conclusion is that using bibliomtric to make analysis on the content of IR is one of new approachs to create and use personas(multi-user model) for expending knowledge service and it maybe helps adding the usage rate of IR.

Keywords Institutional Repository; Bibliometric; Knowledge Service; User Behavior; User Model

移动图书馆解决方案

移动图书馆是指依托成熟的无线移动网络、互联网以及多媒体技术,使人们不受时间、地点和空间的限制,通过各种便携移动设备(手机、PDA、手持阅读器、平板电脑、MP4等)方便灵活地进行图书馆的信息查询、浏览和获取资源内容的服务。目前国内高校数字图书馆要实现上述功能,还存在一些问题,主要是无法解决对馆藏资源的全文访问和对各类数据库资源进行统一检索。要做到这一点并非易事,首先上网资费是一大难题,其次在技术层面又涉及三方面问题,一是大量的数据库平台都不一样,没有通用的接口供访问;二是数据存储的格式不同,各自的加工方式以及加密保护的办法也不一样;三是移动终端品种庞杂,操作系统各不相同。

2010年 6月,北京书生公司推出了"移动图书馆解决方案",有效地解决了上述主要技术难题。首先通过 UOML文档交换服务器,解决了不同数据库平台无法统一访问,以及不同数据存储格式不能通读的问题,其次解决了不同手持设备实现统一搜索的问题,支持塞班(Symbian)、安卓(android)、WindowsMobile、linux、iphone甚至是山寨机等各类型的可上网手机,以及具有 Wi-Fi功能的 MP4、电子书阅读器等手持终端设备,第三突破了过去的 IP控制方式,直接针对每个读者终端进行权限控制。在知识产权保护方面,"书生移动图书馆系统"花大力气进行合法手机用户的惟一性认证,限定每个手机号绑定一次就是一次,如果失效的话就得重新认证。对 MP4以及专用阅读器,也有惟一的 ID绑定。另外对访问过程中链接传播的安全性控制也重点加以考虑。